

S-E-C-R-E-T

#25

Library

21 April 1961

MEMORANDUM FOR: Deputy Director, Intelligence

FROM : Chief, DD/I Automation Staff

SUBJECT : Computer Needs

The DD/I Automation Staff has completed its initial survey of DD/I computer needs. The Staff has investigated numerous "special project" interests in the Research Offices of the DD/I, has participated in the development of Project CHURCHWAY, and has sponsored an intensive six-week survey of automation requirements by teams of specialists from IBM and the System Development Corporation (SDC).

The FINDINGS of this Staff concerning your EDP needs and our RECOMMENDATIONS to you appear below. The following Annexes are also attached: ANNEX A - Summary of IBM and SDC Reports; ANNEX B - The DD/I Computing Requirement; ANNEX C - The Hardware Requirement; and ANNEX D - Estimated Requirements of a DD/I Computer Center.

* * * * *

FINDINGS

A. Principal Findings

1. There is a definite computer need in the DD/I Area. This need may be categorized into two basic types of computer applications:

a. "Special project" applications within the Research Offices involving statistical or other manipulations of data and relating to specific intelligence problems, and

b. Document and information retrieval applications involving your major document collections and relating to the entire spectrum of intelligence interests.

2. The "special project" applications are for the most part entirely suitable to present-day electronic data processing (EDP) capabilities. Data preparation and programming will, of course, be required but extensive system analysis and design are not typically involved. With the establishment of a DD/I computing capability, implementation of these special projects could proceed directly.

S-E-C-R-E-T

3. The document and information retrieval applications will require an extensive system analysis and design effort. This effort should be considered experimental and developmental. Computer support should be available to this effort. The great size of your central reference files and the broad scope of consumer requirements relative thereto pose the prospect of equipment and system requirements which cannot at present be fully satisfied.

4. There is every reason to anticipate a rapid growth in the DD/I computing need. The "special project" applications will certainly expand as capabilities are established and demonstrated. With the design of a successful system for the document and information retrieval function, a very massive workload will result. - There are three additional factors which will foster a growing reliance upon computers in the future:

a. The advanced collection programs for ELINT, RADINT, and COMINT are being designed to acquire information in analog and digital form. With corresponding equipment development on the data-analysis side, most of this information can be made available immediately for computer processing, thereby greatly speeding the realization of intelligence gain from these sources.

b. The growth of computer capabilities within the Intelligence Community will lead to an increase in the exchange among community members of information recorded in machine language formats.

c. The increased use of tape-generating typewriters at Headquarters and in the field will make available in machinable form vast amounts of intelligence data without additional preparation effort.

5. There is a need for basic research in intelligence applications for computer capabilities. Such research would serve to keep the CIA computing effort dynamic and imaginative.

6. The DD/I, and indeed CIA, is very short on the skills required to implement and operate a computing facility. An extensive training effort for Agency personnel will be required. In addition, we must rely for some time to come on outside sources for certain necessary skills.

7. Centralization of EDP Staff functions and of computing facilities will tend to maximize the usefulness of hardware, software (supporting computer programs and operating procedures), and the limited human skills available to us.

8. Years of effort, experience, and policy support are required to achieve an efficient and effective computing capability. Given a vigorous start now, it will still be 1964 or later before the major potential of a DD/I computing capability will have become manifest.

9. Top Management support is absolutely essential to any effort to establish a DD/I computing capability. Such an effort will require money, people, space, etc.; it will also eventually influence operational and organizational structures. The latter consequence must be expected and faced squarely from the start.

10. It should be stressed that the successful prosecution of the extensive system study of the DD/I document and information retrieval function (paragraph 3 above) will require the full Management approval of those Offices affected (OCR is the principal such Office). Experience in this area shows that the best experimental efforts of this kind occur when the personnel and management of the organization concerned are enthusiastic about the objectives and undertake the major share of the work themselves. Without this climate, the wisdom of starting such a laboratory is questionable.

11. The establishment of a computer capability for the DD/I is going to cost money and people. Such a facility proffers not only a better way of doing old things, but also a new facility with new capabilities. Any decision you make to go ahead on the computer front should be based on your basic need for more sophisticated data processing capabilities and not on the hope of saving money.

B. Background Findings

1. Computer capabilities are actively being explored and developed now in every major component of the Intelligence Community. The impact of computers on the structure and operations of the Intelligence Community is already being felt and will grow dramatically in the years ahead.

2. The scope of Community efforts in the computing field today -- and, thereby, some suggestion of the impact for tomorrow -- is indicated by the following list of computers now in use by the Intelligence Community:

NSA ----- Four IBM 704's; three IBM 705's; IBM 650;
UNIVAC 1103; IBM 1401; Harvest on order

Air Force-- Three IBM 7090's; four IBM 1401's; three Minicard
systems

Army ----- Sylvania 9400; CDC 1604 (DOD-DAC)

Navy ----- IBM 1401; IBM 7090

CIA ----- IBM 650 (DD/P); RCA 501 (DD/S); ALWAC (NPIC);
one Minicard system

3. The DD/P has been at work for four years now in designing and implementing an advanced, large-scale document and information retrieval system (Project WALNUT). This system is designed to satisfy internal DD/P needs and will not possess the additional capacity to meet DD/I requirements. Hardware under development for this Project may have applications in the DD/I.

4. The RCA 501 computer of the DD/S is now operational. The DD/I has and will be given some time on this computer although time available to the DD/I will be far from adequate to handle the DD/I workload.

5. NPIC has a small, special purpose computer (ALWAC) for its photogrammetric and other applications. This computer does not constitute a significant EDP capability.

6. The DD/I pioneered the application of machines to intelligence tasks with the establishment a dozen years ago of its OCR punched card system. It has not, however, kept pace with other elements of the Community in second-generation (computer) equipment.

* * * * *

RECOMMENDATIONS

The DD/I Automation Staff recommends that you:

1. Decide now to commit your Area to a major EDP effort and to establish a DD/I Computing Center in the New Building.

2. Create a new operating component, which shall constitute the DD/I Computing Center, to support total DD/I requirements for EDP services.

3. Approve the following short-term objectives for the Center:

a. To apply EDP capabilities to DD/I Special Projects.

b. To perform EDP experimentation with the document and information retrieval requirement.

c. To initiate a broad, EDP-oriented, system design study for optimizing reference capabilities in support of user requirements.

4. Approve the following long-range objectives for the Center:
 - a. To implement and maintain EDP programs devised to support evolving DD/I functions.
 - b. To develop and maintain those human skills required for a dynamic computing program for the DD/I area.
 - c. To monitor advances in computer technology and to plan and provide for the future equipment requirements of the DD/I area as the DD/I computing program develops.
 - d. To sponsor appropriate R & D programs in the application of computer capabilities to intelligence processes.
 - e. To serve as the focal point of representation for the DD/I in EDP matters.
5. Direct the DD/I Computing Center to:
 - a. Commence staffing efforts against a first-year personnel authorization of approximately thirty-five (35) persons.
 - b. Implement the DD/I EDP program in conformance with the short-term objectives in paragraph 3 above.
 - c. Procure the use of outside contractual services to the extent of approximately twenty-five persons for training, system design, programming, etc.
 - d. Issue a letter of intent for the procurement of EDP equipment on a rental basis.
6. Obtain priority support from the DD/S Data Processing Staff to assist in the DD/I planning and implementation effort.



25X1A

ANNEX ASUMMARY OF IBM AND SDC REPORTS

Summarized below are the principal recommendations contained in the recent IBM and SDC surveys. The cost-free IBM and SDC reports were prepared to assist us in the formulation of a sound DD/I plan of action.

1. IBM restricted the scope of its effort to studying OCR functions. SDC, however, was asked to survey, in addition to OCR, the possible computer needs ("special projects") of the DD/I research offices.
2. The IBM PAPER makes suggestions for EDP within BR and SR and recommends a long-range systems analysis and hardware development program. Specifically, IBM finds that:
 - a. OCR punched card systems have reached their optimum effectiveness and the time is at hand to implement conversion to second-generation (EDP) systems;
 - b. A simplified WALNUT document image device offers promise for certain OCR files;
 - c. A WALNUT index device (large capacity, random access, magnetic storage) is applicable to a number of index files in OCR, especially BR;
 - d. SR should be used as a "proving ground" for the development of an advanced OCR-wide data processing system;
 - e. A short-term, cost-free study should be conducted jointly by IBM and CIA on SR and BR to plan the conversion of selected functions from punched cards to magnetic tape. EDP equipment should be procured and conversion effected as indicated by this study;
 - f. A long-term over-all system analysis should be undertaken by competent systems engineering personnel to define detailed requirements for an advanced OCR data processing system.
3. The SDC PAPER stresses comprehensive planning plus top management support as indispensable prerequisites to the introduction of computers within an organization. It says that the need exists now for an intensive systems study aimed at the design of a new information

processing system for the 1965-1970 time scale. SDC devotes considerable attention to the managerial aspects of getting going, and concludes that EDP equipment should be ordered for a DD/I computing facility as soon as possible. Its principal recommendations are that:

- a. "Special project" needs in the research offices represent adequate justification now for the establishment of a DD/I computer facility.
 - b. A part of this facility should be established to perform experimental work on advanced OCR document and information control systems based, at the start, on the files of SR.
 - c. CIA staff up for an Agency data processing program, with outside support, to lay the foundation for an integrated CIA information processing system by 1965.
 - d. Short-range improvements to the existing OCR systems should be minimized.
 - e. A permanent EDP Planning Staff, reporting to the highest feasible administrative level, should be established to advise and assist top management in all areas of information processing in CIA.
 - f. A broad training program should be initiated to develop data processing knowledge and skills in CIA.
 - g. CIA should monitor and support research and development efforts in the EDP field on an increased scale.
 - h. External support should be obtained to assist CIA in (1) the proposed over-all system study, (2) the development of an advanced document and information control system in OCR, (3) the formulation of computer solutions to research office "special project" needs, and (4) the training of CIA personnel.
4. There is considerable contrast between the IBM and SDC papers in the approaches advocated; namely:

IBM	SDC
Suggests a cautious and gradual transition from present methods to EDP.	Stresses that computer capabilities require the structuring of an entirely new system.
Recommends BR and SR convert certain present operations to EDP now.	Warns repeatedly against "patchwork" solutions.
Offers to develop new <u>hardware</u> with CIA support.	Recommends the development of new <u>systems</u> with CIA support.
Pays tribute to the systems competence now on hand in OCR.	Suggests that major portion of skills initially required by CIA must be obtained from external sources.

5. In spite of the divergence suggested by the above, these papers have many fundamental areas of accord between them. For example, both IBM and SDC conclude that:

EDP is definitely applicable to the functions of OCR and the DD/I.
The DD/I should start now to move in the direction of acquiring a strong EDP capability.
The DD/I should decide now to acquire EDP equipment for the DD/I.
A developmental effort should be launched now intended to devise a computer-driven document and information retrieval system suitable for general application within OCR and perhaps other areas of the DD/I.

~~S-E-C-R-E-T~~

Approved For Release 2002/02/06 : CIA-RDP83T00573R000500170003-9

EDP training is of strategic importance to Agency personnel.

An extensive system design effort should be undertaken to achieve optimum advantage of new EDP capabilities by the mid-decade period.

A-4

~~S-E-C-R-E-T~~

Approved For Release 2002/02/06 : CIA-RDP83T00573R000500170003-9

ANNEX B
(Section I)

THE DD/I COMPUTING REQUIREMENT

I. GENERAL

A. The DD/I computing requirement may be categorized into two basic types of computer applications:

1. Special Project applications within the Research Offices involving statistical or other manipulations of data and normally relating to specific intelligence problems.

2. Document and Information retrieval applications involving major document collections and relating to the entire spectrum of intelligence interests.

B. There is every reason to anticipate a rapid growth in the DD/I computing need. The Special Project applications will certainly expand as capabilities are established and demonstrated. With the design of a successful system for the document and information retrieval function, a very massive workload will result.

C. There are additional factors which will foster a growing reliance upon computers in the future:

1. ELINT, RADINT, COMINT, and other collection programs are approaching a capability for rapid, if not simultaneous, conversion of analog data to digitalized forms required for computer processing. Once this capability is achieved, information from these collecting programs can be made available for immediate computer processing, thereby speeding the realization of intelligence gain from these sources.

2. The growth of computer capabilities within the Intelligence Community will lead to an increase in the exchange among community members of information recorded in machine language formats.

3. The increased use of tape-generating typewriters at Headquarters and in the field will make available in machinable form vast amounts of intelligence data without additional preparatory effort.

4. A general up-grading in the level of sophistication in data handling activities is now occurring throughout the Intelligence Community and will certainly continue for some time

B-1

to come. This results in large measure from the introduction of computers into the Community's data processing programs. CIA will desire to and should keep up with -- if not lead -- this development.

D. At the present juncture, the DD/I has no computing capability (with the exception of the small-capacity ALWAC machine in the NPIC). The DD/I is lagging behind virtually all other elements of the U. S. Intelligence Community in the application of computing resources to the intelligence task.

ANNEX B
(Section II)

II. THE SPECIAL PROJECTS REQUIREMENT

A. DEFINITION

1. The term "special projects" as used in this paper refers to individual file and data processing needs which by nature fall outside the normal machine support systems of OCR (hence "special") and which appear suitable for machine handling. "Special projects" are usually characterized by a relatively circumscribed data base of extracted information, by updating requirements, by intensive manipulation of the data base, and by periodic report-generation (print-out of data by machine). Most special projects are located in the DD/I research offices.

2. Special projects constitute one of the two basic types of potential computer applications within the DD/I Area. The DD/I Automation Staff has recommended the establishment of a computing capability to handle the special project need (Recommendation 3a).

B. GROWTH IN ANALYST INTEREST

1. DD/I intelligence analysts frequently encounter file or data processing requirements beyond their own capabilities and not accommodated by the central reference system. This is natural and not necessarily a reflection on the established reference system. It is to be expected that special data processing requirements will always be a part of the DD/I analytical scene... and, indeed, that the number and nature of such requirements will vary directly with the degree to which the DD/I intelligence effort is imaginative, dynamic, and vital.

2. Central to the special project concept is the handling and recall (in forms usable by the intelligence analyst) of information. It is the analyst's need for quick access to information, rather than documents, and, even more important, his frequent need to manipulate such information in diverse and often sophisticated fashions which has generated the expanding interest in special project services. There is every reason to anticipate a continued growth - if not, indeed, an explosion - in the demand for machine-driven research support of this nature.

B-3

S-E-C-R-E-T

S-E-C-R-E-T

3. There are many factors fostering development of the DD/I interest in special projects. Some of the more basic are:

a. The Agency's document collections have grown to such sizes it is often impossible for the research analyst to exploit the large number of documents provided to him in response to his requests.

b. The period of exhaustive, basic research within the DD/I has largely come to an end. The ponderous 200-300 page study, based on detailed analysis of all available materials, is no longer being written. The DD/I analyst's reports are now much briefer, their topics more circumscribed, and their chronological span more restricted. The DD/I analyst is now endeavoring to update what has previously been published... to report changes in status. He is increasingly concerned with current reporting. And he is often working against very short deadlines. He has less need and less time for methodical perusing of documents.

c. Data analysis techniques employed in the DD/I are becoming more sophisticated. Statistical analysis, extrapolation and interpolation, model building, etc. are assuming increased importance to the DD/I analyst in his efforts to produce better intelligence. Such techniques are often difficult or impossible to use without machine support. The availability of a computing capability within the DD/I will give further impetus to the use of such techniques.

d. The DD/I is now placing heavy emphasis on special weapons intelligence. The intelligence effort in this field often requires the sifting of large quantities of relatively low grade materials in order to extract pertinent fragments of data. These fragments must often then be sorted, correlated, or otherwise manipulated in the effort to generate an intelligence picture. At the same time, the analyst is faced with an increasingly urgent demand for timeliness in his reporting. It is only natural that, faced with these requirements, the special weapons analyst has turned toward the machine for support.

e. In several other parts of the Intelligence Community, computing resources are today being applied to data problems generically similar to those faced by our DD/I analysts. Our analysts are, of course, exposed to these computing efforts through contact with their counterparts in other organizations.

3-4

S-E-C-R-E-T

Their exposure to these efforts has done much to spark a desire within our DD/I Offices to bring similar resources to bear upon their own problems.

C. GROWTH IN MACHINE CAPABILITIES

1. While the need for special project support has been growing, the capabilities of machines to assist in this type of effort have been dramatically extended. With the advent of the computer, an entire new dimension in the storage, handling, analysis, and retrieval of information by machine has resulted. Computers provide a greatly increased power to store in machinable form large masses of information, to search out and locate desired information, to manipulate data in intricate and logically sophisticated fashions, and to print out information in natural language and with full command over the editorial formats of print-out. All of these functions are performed, of course, at very high speeds. The appreciation of the above on the part of the DD/I Offices has naturally served to stimulate further their interest in the special project aspect of the DD/I data processing outlook.

D. TYPES AND EXAMPLES OF SPECIAL PROJECTS

1. The DD/I Automation Staff has identified several potential special projects in the DD/I Area. There is considerable variation among these projects in terms of data volumes, nature and complexity of data manipulation and output requirements, ease of implementing for computer processing, importance of the intelligence problem to which they relate, prospect of intelligence gain, etc. These projects do not lend themselves to clear-cut and distinct categorization. It may be helpful, nonetheless, to indicate three major types of applications represented with examples of each:

a. File Maintenance and Report Generation

This type of special project application is characterized by medium-to-large input/output volumes and minimal internal processing requirements. Such projects are relatively easily planned and activated. Programming requirements are standard and often satisfied by library routines supplied by the equipment manufacturer. Large machine memory is not normally necessary but strong input/output and peripheral equipment capabilities are required. Preparation of input data for machine processing

B-5

S-E-C-R-E-T

is typically a sizeable effort.

Examples:



25X1B

b. Mathematical and Statistical Analysis

Input/output volumes are, in general, more moderate and internal processing requirements greater for this type of application than is the case with file maintenance tasks. Most of the mathematical techniques involved are well known and established; some are more advanced, however, such as numerical approximations to dynamic models. The formulation stage is considerably more complex for this category of applications. Extensive consultation with the sponsoring intelligence analysts will be necessary and greater analytic capability on the parts of the formulators and the computer programmers is required. The programming requirement is typically complex. Large amounts of machine memory may be required.



25X1B

c. Advanced Systems

Applications falling into this category will tend to be large-scale computing systems of complex design, using advanced hardware and advanced techniques. These applications should be considered in large part developmental. All aspects of data handling and computing may be involved, from the most simple to the most challenging. These projects or systems will be characterized by such aspects of computing as: pattern recognition and analysis, hypothesis testing, machine "learning", automatic syntactic analysis of input data, advanced visual displays, remote inquiry

B-6

S-E-C-R-E-T

S-E-C-R-E-T

and man/machine interaction, etc. This type of application stresses creative participation in the system by the intelligence analyst and taxes the automation potential to the utmost. System design, computing procedure formulation, programming, hardware configuration, etc. will be expected to progress through several stages of development and interplay before maximization of the project is achieved.

Examples:

- (1) Project CHURCHWAY
- (2) Project SNOWFLOW (CIA Counterpart)
- (3) National Indications Project

E. PROSPECT

1. The data-processing requirements of special projects are, for the most part, entirely compatible with present-day EDP capabilities. Data preparation and programming will, of course, be required but extensive system analysis and design are not typically involved. Much of the work of implementing selected projects can be carried out prior to the installation of computing hardware. Analysis of input data and project objectives, project design, preparation of input data for machine processing, writing of computer programs, etc. can start as soon as manpower becomes available. De-bugging of programs written during this period can be carried out on equipment available elsewhere in the Washington area. Thus an operational capability for special projects should be a reality directly subsequent to installation of hardware and engineering checkout.

2. At present, the DD/I is not organized to handle the special project need of its analysts. To the extent this need has been met, it has been met by superimposing the special project workload upon organizations staffed and equipped for other activities. The Special Register in OCR has been the most aggressive component within the DD/I in the special projects area and has done much with its punched-card equipment and limited manpower resources. But support to this need falls far short of what is required for a dynamic and advanced special projects capability for the DD/I. The Special Project need can be adequately provided for only by establishing an organizational element charged with meeting this need and equipping that component with EDP capabilities.

B-7

S-E-C-R-E-T

ANNEX B
(Section III)

III. THE DOCUMENT AND INFORMATION RETRIEVAL REQUIREMENT

Before computer capabilities can be applied to the DD/I document and information retrieval function, a comprehensive system design effort must be conducted. The objective of this effort will be the structuring of an advanced, EDP-oriented, document and information retrieval system which will satisfy the real reference needs of the DD/I for the 1960 decade. This system design project should be computer-supported. (Recommendations 3b and 3c)

* * * *

A. BACKGROUND

1. The system design undertaking recommended by the DD/I Automation Staff will be DD/I - wide in scope. The organizational component most directly concerned, however, will be OCR... where the document and information retrieval function is largely centralized.

2. The OCR reference system is unique in the scope and complexity of its task. From the earliest days OCR management has sought to exploit the potential of machines to assist in this gigantic document and information control effort. OCR has, over the years, developed machine indexing skills in the area of its responsibility which are unequalled elsewhere in the Community.

3. OCR is now operating a series of very large-volume, punched-card index systems. Each of these systems has its strengths and each has its problem areas. It is easy and perhaps natural for the user of OCR services to find fault with the reference responses received from these systems. Unfortunately, however, many users' criticisms are parochial in nature and do not pave the way to real and substantial improvements in the OCR system.

B. PROBLEM AREAS

1. The above is not to say that real difficulties do not exist in the OCR system, or that our reference services cannot be improved. There are several problem areas of concern to OCR management as well as user components. Many of these problem

areas have been examined and commented on by past survey groups and it would be redundant here to dwell on them. DD/I officials and analysts are well aware of such problems as: multiple service points within OCR; imprecise document retrieval; "wheelbarrow" volume of system responses; incomplete index coverage; ad hoc processing techniques which compromise system integrity; duplicative input processing among OCR components and, more serious, redundancy in documents furnished by these multiple service points to a given consumer on a given problem; inadequate capability for information (vs. document) retrieval; inability of the system to respond effectively to certain new or specialized recovery requirements; limitations on the length of the machine record and therefore the amount and form of data recorded; coded information output instead of natural language output; space requirements, etc.

2. Some of the above problems have reached such serious proportions as to endanger certain of the fundamental objectives of the OCR service system. Efforts to attack these problems have been made largely within the context and constraints of an OCR essentially unchanged since its original establishment. Although punched-card speeds and capabilities have increased over the years and OCR has made certain system improvements - some of considerable significance -, the fact remains that the overall reference effort is operating today in terms of the basic organizational, schematic, and equipment characteristics of a decade ago.

C. PROGRAM

1. It is time for a serious, intensive, and comprehensive study of the DD/I document and information retrieval function and for a major effort to design a new and advanced document and information retrieval system. It is time because:

- a. The OCR reference system has not changed basically since its inception.
- b. Real problems exist today.
- c. The solutions to these problems do not appear to lie within the tolerances of the present OCR system. OCR has worked diligently to maximize the effectiveness of its operations and it therefore seems unlikely that new and major extensions of OCR's capabilities can be realized within our present systematic, organizational, and hardware posture.

- d. Second-generation equipments (computers) are now available which offer dramatic new data processing capabilities. The new dimensions in hardware capabilities argue for a design of new dimensions in system capabilities.
- e. Advances with computer applications elsewhere in the Intelligence Community are such that the central position of OCR in its field of responsibility may, indeed, soon be placed in jeopardy without such a study and design effort.

2. Let it be squarely faced, however, that EDP offers no guarantee of success. EDP is expensive. It is complex. Its application in the document and information retrieval field is still largely uncharted. EDP will place new and exacting demands on the skill and ingenuity of our personnel and on the structure and integrity of the data processing systems we evolve. Experience has demonstrated, however, that machines are essential to the sophisticated control of large-volume document and information collections. An extensive system design effort, embodying an upgrading of machine support from first-generation (punched card) to second-generation (computer) capabilities... and building on hardware capabilities which are today only anticipated... does appear to constitute the logical next step in the evolution of our DD/I reference system and, as such, offers the best prospect of success.

3. The DD/I Automation Staff has recommended a broad, EDP-oriented, system design effort (Recommendation 3c) and an EDP experimentation program (Recommendation 3b) for the document and information retrieval function to be conducted by CIA and contractor personnel.

- a. The basic objectives of the system design effort are:
 - (1) To anticipate and define the proper, consumer-oriented, goals of our reference service for the 1960 decade.
 - (2) To devise an advanced, computer-driven system to fulfill those goals.
 - (3) To evolve methods and techniques for performing, on an integrated basis, the constituent functions of the system.

B-10

S-E-C-R-E-T

- (4) To determine the informational, processing, personnel, and equipment requirements of the system.
- (5) To construct an over-all implementation plan for the system, - including, where applicable, conversion, parallel operation, and phase-out of present systems.

b. The basic objectives of the experimentation program are:

- (1) To provide a laboratory for operational development and computer testing of concepts and techniques surfaced by the system design effort.

Certain of these experimental areas can be anticipated; e.g., new index languages, based on normal English words, to facilitate data input and permit direct output from the machine in natural language form; table look-up applications to permit automatic conversion from natural language to digital code and vice versa; automatic indexing to permit the indexing by machine of information available in machinable form; computer testing of divers schemes of file organization; automatic dissemination; and, perhaps, automatic abstracting.

- (2) To plan and carry out an experimental conversion operation for the Special Register system.

This conversion operation will be based on present SR input indexing. Input data will be restructured, however, to take advantage of EDP record lengths and search capabilities. Both serial sequence (tape reels) and direct or random access index storage techniques will be simulated and evaluated.

- (3) To plan and test on the computer an experimental model of an all-source, all-topic, document and information retrieval system.

This model will serve to test the feasibility

B-11

S-E-C-R-E-T

and capabilities of a single, integrated, OCR-wide document and information retrieval system.

(4) To plan and test models of systems structured by topic rather than source document (e.g., a biographic information system), and other such models as is appropriate.

(5) To experiment with methods of computerizing selected sub-systems within the OCR Registers as seems profitable.

(6) To foster the development within the DD/I of computer-related human skills in the document and information retrieval field.

4. An important segment of both the system design and experimentation programs will be the structuring and testing of a single, all-topic system for document and information retrieval which will provide for:

a. A single reference point for OCR, regardless of topic on which reference support is desired.

The customer would place one integrated request which would activate one integrated servicing effort which would, in turn, produce one integrated service product at one point in time. The desirability of such a reference support structure has repeatedly been expressed by OCR consumers.

b. The elimination or reduction of multiple input processing by providing for a single input effort to control all elements of a given document.

The multiple input processing which sometimes occurs within the present OCR system consumes manpower and thereby contributes to the processing backlogs and ad hoc filing and control arrangements which have taken on such serious proportions.

c. An improved "pertinency ratio" for documents or information recovered from the system.

B-12

S-E-C-R-E-T

An all-topic control within a single, homogeneous system, preserving valid data relationships and providing for flexible recovery by logical combinations of data, would enhance both the subtlety and the specificity of document and information recovery.

d. The expansion of information (as distinguished from document) retrieval capabilities of the OCR machine system.

An increased information retrieval capability would speed the research analyst more directly to the data he needs and would reduce the document retrieval requirement. This is a particularly fertile area for development.

B-13

S-E-C-R-E-T

ANNEX C

THE HARDWARE REQUIREMENT

I. LETTER OF INTENT

The DD/I Automation Staff has recommended that a letter of intent be issued as soon as possible for the delivery in approximately twelve months of EDP equipment on a rental basis for the DD/I Computing Center (Recommendation 5d).

II. GENERAL CONSIDERATIONS IN SELECTING EQUIPMENT

The following general considerations bear upon the selection of computing equipment:

A. The equipment should be well-proven to insure reliability.

This will promote maximum utilization of available machine time and minimize re-runs due to machine malfunction.

B. The equipment should be modular in design.

"Modular" refers to the building-block approach in the design of equipment. Such design contributes to ease and speed of maintenance and permits alteration of system capabilities.

C. Equipment which is widely accepted and used should be selected.

Such equipment will carry with it a strong "software" capability. ("Software" is the supporting programming and system documentation necessary for efficient utilization of any EDP complex. It includes manuals, service routines, high-level programming languages, and a wide variety of operating procedures. The ready availability of required software can expedite and greatly reduce the cost of operations.)

~~S-E-C-R-E-T~~

D. The equipment should be able to accept as inputs the machine language products from other data processing computer systems in the Intelligence Community.

This means either the equipment should be of the same family as the majority of other systems or that it must have the capability to convert products and programs from these other systems into a form acceptable to the equipment.

E. Equipments acquired should be from the same family (generally this means the same manufacturer).

This assures functional compatibility within the computer complex - an important consideration from both the hardware and software points of view.

F. The family of equipment should include a wide variety of modules or elements, both in type and capability.

This increases flexibility in facilitating interchange and regrouping to accommodate changes in processing or organizational requirements.

G. The chosen manufacturer should have extensive research capabilities for continued developments of their product line.

This allows the user to plan for more advanced capabilities in the future with some confidence that new equipment developments will proceed apace.

III. DD/I CONSIDERATIONS

A. Computing applications presently identified within the DD/I fall into two basic types: special project applications and document and information retrieval applications. Most medium to large scale computer systems could handle both types of requirements. Certain of the special projects to be activated, however, are better suited to implementation on a scientific computer. The document and information retrieval applications and certain other special projects are better suited to implementation on a data processing computer.

C-2

S E C R E T

The term "scientific" computer is generally used to denote a computer system having extremely powerful arithmetic capabilities built into its design - such as fixed word length memory and parallel transfer. A system of such design is better suited to performing mathematical computations than to manipulating large volumes and varieties of alpha/numeric and natural language data. The "data processing" computer, on the other hand, is engineered to handle well these varied alpha/numeric and natural language data but is less suited to performing mathematical computations.

B. The above suggests acquisition of two computers for the DD/I Computer Center: a scientific computer and a data processing computer.

C. It is the opinion of the DD/I Automation Staff, however, that the initial letter of intent should provide for one computer only. It is also the Staff's opinion that, in view of the over-all nature of the DD/I requirement, the computer ordered now should be in the data processing class rather than a scientific computer.

D. In addition to the general selection considerations in paragraph II above, the following specifics should be looked for in selecting the initial computer system for the DD/I:

1. Flexible and powerful input/output capabilities (accepting punched cards, paper tape, and magnetic tape, and having high speed printer)
2. Flexible and efficient data manipulation capabilities within the main frame (variable word length and character access)
3. Large random access storage capacity
4. Numerous (15/) magnetic tape drives
5. High speed, high density magnetic tape characteristics
6. Medium size core memory

C-3

S-E-C-R-E-T

7. Extensive buffering
8. Extensive instruction repertoire
9. Adequate complement of registers
10. Parallel processing/time sharing features
11. Good interrupt features
12. Medium arithmetic capability
13. Effective console control
14. Typewriter inquiry station
15. On line/off line switching capability

E. The Staff's conclusion that only one machine should be ordered at this time is based on the judgment that additions to the hardware complex of the DD/I Computer Center should be considered after we have gained more experience in the DD/I effort and after the direction of the proposed system design effort on the DD/I document and information retrieval function has become manifest. The expected availability of time on the DD/S RCA 501 computer for DD/I applications has constituted a significant factor in reaching this conclusion.

F. With the expansion of the DD/I special project effort, the prospective need for a scientific computer must be recognized. The development of a computer-driven system for the DD/I document and information retrieval function will clearly generate major new hardware requirements. CIA management should be aware of this outlook.

C-4

S-E-C-R-E-T

25X1A

Approved For Release 2002/02/06 : CIA-RDP83T00573R000500170003-9

Approved For Release 2002/02/06 : CIA-RDP83T00573R000500170003-9

ITEM #2: Contractor Personnel

A contract should be let with a suitable outside organization to provide needed skills not available within the Agency. The main aim of the contract will be the redesign and structuring of an EDP-oriented information processing system for the DD/I. Approximately twenty-five (25) people should be so engaged at the start. The requirement of an additional ten to fifteen people is anticipated for the second year. Skills provided should include: training specialists, system analysts, programmers, and EDP managerial specialists. This group would be expected to reach its greatest strength during the second year. The contractor group should taper off to the zero point by the end of the fourth year. This group will work closely with CIA personnel and in support of CIA personnel.

ITEM #3: Equipment Rental

A letter of intent should be placed as soon as possible for the rental of suitable EDP equipment. Installation of this equipment will take place during FY 63; therefore, no funds for this purpose will be required during FY 62.

Punched card and punched paper tape generating equipment will be required from the start to support the preparation of input data and computer programs for special projects and experimental activities on the document and information retrieval effort. Two card punch machines and two card verifier machines should be rented; it will probably prove advantageous, however, to purchase the paper tape generating equipment.

ITEM #4: Supplies and Equipment

Supplies - Supplies to support the computer center will include: paper tapes, magnetic tapes, punch cards, continuous form printer paper, operational control forms, etc.

Equipment - Two paper tape punch-verifier machines should be purchased to support the Special Projects and Document and Information retrieval programs. A variety of housekeeping items (card and tape file cabinets, special work tables and trucks, paper separating machines, etc.) will be needed.

ITEM #5: Programming Support

Programming capabilities will be developed within the CIA staff. The contractor organization will supply programming skills, as well. It is anticipated, however, that programming requirements will occur, particularly during the early years

of this EDP effort, which will best be met under purchase order or service contract with an outside organization.

ITEM #6: Outside Training

Two basic types of training will be needed: Technical Training and Personnel Orientation Training. Most of the training needed will be given in the Washington area by CIA personnel, contractor personnel, and/or hardware manufacturer(s).

Some participation, however, in training programs offered by other government agencies and academic and professional organizations will be required. ITEM #6 is to cover the costs of such training.

ITEM #7: Travel

Included in this item are travel expenses to visit contractor organizations, equipment manufacturers, out-of-town systems of interest to the DD/I, conventions, etc.

ITEM #8: R & D Effort

An R & D effort in areas of interest to intelligence should be sponsored by the Agency. This effort should concern techniques and systems primarily, rather than hardware, and should complement and support developmental activities undertaken within the DD/I in establishing its computing program.

ITEM #9: Site Preparation

Provision for requisite electric power, air conditioning, and floor loads will constitute a modest site preparation effort in the New Building. Specific requirements will be provided by the hardware manufacture after firm selection of equipment has been made.

ITEM #10: Space

The space requirements for the Computer Center will approximate the following:

Computers	3,000 sq. ft.
Office Space for Management, Programming and System Personnel	4,000 sq. ft.
Punch Room	1,000 sq. ft.

25X1A

Approved For Release 2002/02/06 : CIA-RDP83T00573R000500170003-9

Approved For Release 2002/02/06 : CIA-RDP83T00573R000500170003-9

March 21, 1962

STATINTL

[REDACTED]
2430 E Street, N.W.
Washington 25, D.C.

STATINTL

[REDACTED]
We are pleased to furnish the recommended schedule for installing the IBM 1401 Data Processing System to provide faster, more accurate input/output data to the installed computer system, as outlined in my letter [REDACTED] on March 13, 1962.

STATINTL

In June 1961 we started a study of the peripheral requirements to the installed computer system as well as the applications being performed on unit record equipment. Following that study a proposal for an IBM 1401 system was presented in September 1961. Close association has been maintained in these areas and in February 1962 a more detailed study was made of the I/O requirements. We looked into the actual record formats on magnetic tape and the report formats required for all the applications currently processed on the installed computer system.

In preparing this schedule, careful consideration was given to minimizing the disruption of current operations, as well as maximizing the efficient installation of your 1401. Based on the knowledge we have gained of the requirements in this area, we are confident that the schedule which we have developed represents the most realistic approach to achieving these prime goals.

A joint effort between your staff and IBM is your assurance of a smooth transition and completely successful solution to your I/O requirements. There is no substitute for experience. Your people have this experience in their detailed knowledge of the applications and IBM has it from the many successful years in the data processing field. This combination will result in the most successful installation of processing these I/O functions.

- 2 -

We have made every effort to insure the inclusion of all information required for the planning and control of a successful operation of this system. We will be happy to provide any additional information you may require. We would like to express our appreciation to you and everyone who assisted in providing the information pertinent to this report.

As your IBM representative, let me assure you that I will personally exercise close coordination on all matters relating to a successful installation of your IBM 1401. In addition, an IBM Systems Engineer will be assigned to work with your people in following the schedule outlined in this report.

We are prepared to start immediately toward the goal of a successful input/output operation.

Very truly yours,



R. D. Pease
Marketing Representative

IBM

Data Processing

General Approach for Conversion

After a detailed study of the work currently performed by the equipment peripheral to the RCA 501, IBM recommends a general approach to converting to a 1401 as a solution to your current input/output requirements. We are confident that this approach will assure a smooth transition from the current plugboard equipment to a stored program system, with a minimum of disruption of current operations.

Currently, there are some 600-700 RCA 501 programs involving a certain amount of card-to-tape, tape-to-card, and tape-to-printer operations. These peripheral operations involve the use of approximately 60 plugboards.

Some of the plugboards, obviously, are used by more than one 501 program. Rather than write an individual 1401 program for each of the 600-700 RCA 501 programs which require peripheral support, we propose that a 1401 program be written to simulate each of the plugboards currently used on the RCA peripheral equipment.

There is a great deal of logical similarity among plugboards, the main differences being the record and report formats. Due to this similarity, we recommend that



Data Processing

a representative program be written for the three peripheral functions; that is, card-to-tape, tape-to-card, and tape-to-printer. These programs could be written, under the supervision of the IBM Systems Engineer, as part of the case problems in the 1401 programming course that will be conducted for your personnel. Therefore, with the general logic established in these programs it will apply to all required 1401 programs. In fact, whole routines could be lifted from these model programs and applied by your personnel to the specific jobs to which they are assigned. In this connection, we recommend that, wherever possible, your programmers be assigned 1401 peripheral programs which relate directly to the RCA 501 programs for which they are responsible. We believe that this plugboard simulation approach will assure a smoother, more logical conversion to the 1401 in the shortest possible time.

IBM's experience of installing data processing equipment has proven to us that a successful operation requires a proper schedule. We have developed a suggested time schedule for the successful installation of an IBM 1401 to process the input/output operations to the RCA 501.



Data Processing

SUGGESTED TIME SCHEDULE

Action To Be Taken

Weeks Prior To Installation

	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
A. Current Work																
B. Initial Problem Definitions	—															
C. RCA 581 Delivery	—															
D. Physical Installation Planning	—															
E. Education	—															
F. Block Diagram and Programming	—															
G. First Test	—															
H. Test Data Preparation	—															
I. Program Assembly and Test	—															
J. Accessories and Supplies	—															
K. Finalize Installation Planning	—															

NOTE: The following pages give a brief explanation of each action.

IBM

Data Processing

A. Current Work

IBM recognizes that the everyday work has to continue when making a conversion from your present input/output operations to the IBM 1401. It is understood that the first week of each month requires full time of your staff of programmers for current work. Obviously, there can always be unexpected emergencies arising during the month that will also require the time of these programmers.

The schedule for the implementation of the IBM 1401 to perform the input/output operations for the RCA 501 has been set up to allow for a minimum of uninterrupted operation during the conversion. This will be achieved by a coordinated effort between your staff and a highly trained IBM Systems Engineer.

B. Initial Problem Definitions

The first week of the conversion period will be used to design the optimum approach that should be taken in programming the IBM 1401 for the input/output operations. By simulating the control panels now used for card-to-tape, tape-to-card, and tape-to-printer, the easiest conversion with the least amount of programming can be accomplished. A basic program design will be established and can be followed for all programs for the 1401, since there is a great deal of similarity in control panels. The IBM Systems Engineer will assist in the design of the basic programming approach.



Data Processing

Further study may reveal that a generalized program can be written to simulate a number of control panels which will greatly reduce the programming effort required for the conversion.

C. RCA 581 Delivery

The great success which IBM has experienced in the installation of data processing systems is due to the tests to which its systems are subjected before being released to the customer. An important part of this test procedure provides for a thorough test of the entire system as a working system, apart from the testing of each component of equipment. In view of this, an RCA 581 Tape Station is required to complete such test on the IBM 1401. It is strongly recommended that the RCA Tape Station which will be used as part of your IBM 1401 system, be used in this test. We will coordinate with you the time and destination for shipment of this tape station, so as to arrive at the IBM plant during the first week of the preinstallation period.

D. Physical Installation Planning

IBM has Physical Installation Engineers that are highly trained in floor space, power and air conditioning as relates to 1401 installation requirements. They will meet with your people during the first week of the schedule to plan for the optimum layout of the 1401 and to insure the adequacy of any other considerations related to physical installation.



Data Processing

E. Education

Past experience has proven that the most successful programs have been written by people who are intimately familiar with the job and have been trained in programming. Our recommendation is to train the six programmers now assigned to the input/output programs for the RCA 501 system.

An on site 1401 programming course will be conducted by IBM. The course will be "tailor-made" to fit your requirements. In view of the requirement to maintain the current work load, this course will be scheduled on a half-day basis. The course will include a field trip to familiarize the students with the IBM 1401.

The 1401 programming course will benefit these programmers in the future. As the 501 programs are modified to take advantage of the flexibility in the 1401, such as page heading, tape controlled carriage for spacing, and page numbering, the 1401 programs can also be modified to insure maximum benefit from this widely accepted Data Processing System.

We also recommend that enrollments in the 1401 Executive Seminars be made during this period for management associated with Data Processing. Arrangements for such seminars will be coordinated with your staff.

F. Block Diagram and Program

Following the IBM 1401 Programming course, block diagrams and programs will be prepared by the programming staff to simulate each of the



Data Processing

control panels relating to the input and output functions of existing programs. Your programming staff will be assisted by our IBM Systems Engineer who has been trained to a high degree of proficiency in assisting and guiding customers in the successful installation of data processing systems.

G. First Test

Experience has shown that it is most important to schedule an early initial test session. The experience which your programming staff will gain from the opportunity to get "hands-on" training on a 1401, as well as to be able to evaluate the results from the actual performance of these initial programs will prove to be invaluable to the successful completion of the total programming effort.

H. Test Data Preparation

Experience has proven the importance of accurate and representative test data. It is necessary to develop data that will test all conditions that occur in a program. For test purposes, it is possible to simulate data that is on RCA tapes and put it on IBM tapes. This will enable you to utilize the 1401 currently installed in the Agency for program testing. Also, the Washington IBM Test Center will be available as an alternate test site. Test data developed on RCA tapes can be tested at Social Security in Baltimore.

I. Program Assembly and Test

Since programming in machine language is more difficult and time consuming, the programming staff will be trained to use a symbolic language

best suited to your applications. These symbolic programs will be converted automatically to machine language prior to program testing. As the applications are programmed and the test data is developed, program test sessions can be conducted.

The 1401 currently installed in the Agency can be used for this purpose. In the event that this cannot be scheduled, IBM will provide adequate time for such conversion, utilizing the 1401 System at the Washington Test Center.

Social Security in Baltimore, currently has a 1401 installed performing many of the same peripheral functions, in support of magnetic tapes created by an RCA 501, similar to your operations. IBM will assist in arranging for adequate testing time on this system. The IBM Systems Engineer will accompany your personnel on such testing sessions, so as to be of assistance during the test, as well as to be able to evaluate the test results and advise management of progress being made.

J. Accessories and Supplies

Your IBM Marketing Representative will meet with your planning staff to insure that full consideration is given to accessories and supplies required in support of your 1401 System. Cards and report forms currently being used are capable of being used on the 1401 System. However, consideration should be given to the use of standard quality paper for the 1401 as soon as the supply of special stock paper acquired for the RCA 501 has been exhausted. In addition, action should be taken at this time to order a sufficient supply of 1403 printer ribbons. Your

IBM Systems Engineer has been trained to assist your personnel in estimating the

Approved For Release 2002/02/06 : CIA-RDP83T00573R000500170003-9

quantity of such ribbons which should be kept on hand. Experience gained at the 1401 Installation in your Agency should also be great assistance in such estimates. A pad of tapes for the tape controlled carriage also should be obtained. IBM will provide a tape punch at no cost to your organization.

K. Finalize Installation Planning

During this period, a final review will be made on all matters relating to the installation of your IBM 1401 system. Both your IBM representative and the IBM Physical Planning Engineer will coordinate with your key personnel on detailed arrangements for bringing the system into the building (e.g. availability of rigging equipment, entrance clearances, etc.). A final review will also be made as to the location of the 1401 system relative to the RCA 501, so as to insure proper cable lengths and adequate sources of power. While the power and air conditioning currently in the installation appear to be more than adequate, this, too, will be confirmed, so as to leave no room for unanticipated errors. IBM will follow on the shipping details with the IBM Traffic Department and provide close liaison with your key personnel as to the date and time of the system arrival. Finally, your IBM representative will insure the availability of a highly trained Customer Engineer to provide the highest standard of maintenance for your system.

Some of the advantages that you will realize with the installation of the IBM 1401 Data Processing System to process your input/output functions are:

1. Delays and re-runs of your off-line reports will be reduced by the proven high reliability of the 1401.
2. Saving additional cost of special forms will be realized through the ability of the IBM 1401 to produce up to eight legible copies of high quality printing using standard paper and carbons.
3. Reduced processing time and improved tape packing on the 501 is possible because the 1401 can generate information such as heading lines, classification lines, page counts, etc.
4. Elimination of a run currently required on the 501 can be accomplished by the 1401 through editing and formatting records for punching cards off-line.
5. Cost reduction of programming, as well as simplifying and expediting program preparation of your input/output functions, is realized through the complete programming systems packages furnished by IBM.